ABSTRACT OF THE DISCLOSURE

A method for converting text to concatenated voice by utilizing a digital voice library and a set of playback rules is provided. Multiple voice recordings correspond to a single speech item and represent various inflections of that single speech item. The method includes determining syllable count and impact value for each speech item in a sequence of speech items. A desired inflection for each speech item is determined based on the syllable count and the impact value and further based on a set of playback rules. A sequence of voice recordings is determined by determining a voice recording for each speech item based on the desired inflection and based on the available voice recordings that correspond to the particular speech item. Voice data are generated based on a sequence of voice recordings by concatenating adjacent recordings in the sequence of voice recordings.